HHS Public Access

Author manuscript

Nat Microbiol. Author manuscript; available in PMC 2020 August 24.

Published in final edited form as:

Nat Microbiol. 2020 April ; 5(4): 545–553. doi:10.1038/s41564-020-0667-3.

FttA is a CPSF73 homologue that terminates transcription in Archaea

Travis J. Sanders[1], **Breanna R. Wenck**[1], **Jocelyn N. Selan**[1], **Mathew P. Barker**[1], **Stavros A. Trimmer**[1], **Julie E. Walker**[1,2], **Thomas J. Santangelo**[1,*]

[1]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523, USA

[2]Present address: Watchmaker Genomics, Boulder, Colorado 80303, USA

Abstract

Regulated gene expression is achieved in large part by controlling the activities of essential, multi-subunit RNA polymerase transcription elongation complexes (TECs). The extreme stability required of TECs to processively transcribe large genomic regions necessitates robust mechanisms to terminate transcription. Efficient transcription termination is particularly critical for gene-dense bacterial and archaeal genomes[1-3] wherein continued transcription would necessarily transcribe immediately adjacent genes, result in conflicts between the transcription and replication apparatuses[4-6] and the coupling of transcription and translation[7,8] would permit loading of ribosomes onto aberrant transcripts. Only select sequences or transcription termination factors can disrupt the otherwise extremely stable TEC and we demonstrate that one of the last universally conserved archaeal proteins with unknown biological function is the Factor that terminates transcription in Archaea (FttA). FttA resolves the dichotomy of a prokaryotic gene structure (operons and polarity) and eukaryotic molecular homology (general transcription apparatus) observed in Archaea. This missing-link between prokaryotic and eukaryotic transcription regulation provides the most parsimonious link to the evolution of the processing activities involved in RNA 3'-end formation in Eukarya.

Transcription termination (Extended Data Figure 1), driven either by DNA sequence and encoded RNA structures (*e.g.* intrinsic termination) or by protein factors (*e.g.* factor-

*Correspondence to: thomas.santangelo@colostate.edu.

Author Contributions

T.J. Sanders performed *in vitro* transcription, RNase digestions, designed and purified FttA variants, designed and assisted in construction of *T. kodakarensis* strain IR5, purified proteins and templates, generated qRT-PCR data, analyzed data, built structural models, performed Western blots, prepared figures and helped write the manuscript. B.R.W. performed *in vitro* transcription, RNase digestions, designed templates and FttA variants, manipulated TK1428 genomic sequences and prepared figures. J.N.S. generated FttA variants, purified transcription proteins, performed bioinformatic analyses, manipulated TK1428 genomic sequences to generate strain TK1428D, built structural models, and prepared figures. M.P.B. and S.A.T. manipulated TK1428 genomic sequences, generated and analyzed qRT-PCR data and assisted with Western blots and protein purifications. J.E.W. purified FttA variants and performed *in vitro* transcription. T.J. Santangelo conceived and directed the project, wrote the manuscript, analyzed data, and prepared figures.

Reporting summary. Further information on the research design is available in the Nature Research Reporting Summary linked to this article.

Competing Interests

The authors declare no competing interests.

dependent termination) ensures rapid dissociation of RNA polymerase (RNAP) from the DNA template to recycle RNAP and generate RNA 3' ends[1,9]. While often prevalent within prokaryotic genomes, intrinsic termination sequences are typically neither sufficiently abundant nor efficient to mediate all termination events. Transcription termination factors must then efficiently recognize TECs that are not intrinsically terminated and compete with continued elongation to mediate release of the nascent transcript. While the identification of Eta provided evidence of factor-dependent archaeal termination[10], no kinetically-efficient mechanism of factor-dependent archaeal transcription termination has been described. The retention of operon-organized archaeal genomes and the sensitivity of the archaeal transcription apparatus to bacterial rho-mediated termination *in vitro*[2] – combined with the normal coupling of transcription and translation[8], the resultant polar suppression of downstream expression in the absence of such coupling in archaeal cells[11] and the conservation of Spt5/NusG in all genomes – implied the existence of a kinetically-relevant archaeal transcription termination activity that might function akin to the bacterial rho protein. Rho homologues are, however, restricted to Bacteria[12], arguing instead that conserved archaeal-eukaryotic or unique archaeal factors may drive factor-dependent archaeal transcription termination.

Only a core set of ~200 gene families (more properly, archaeal clusters of orthologous genes; arCOGs) are conserved in most archaeal genomes [13], just ~129 arCOGs are strictly ubiquitous[14] and one is an obvious orthologue of a subunit of the cleavage and polyadenylation specificity factor (CPSF) complex [13-18]. The homology of most archaeal transcription components to eukaryotic factors argued that the archaeal homologue of eukaryotic CPSF73 might function as the Factor that terminates transcription in Archaea (FttA). We challenged promoter-initiated TECs [19,20] – generated with an RNAP variant with a His_6/HA-epitope-tagged RpoL subunit [19,21,22] and containing a radiolabeled nascent transcript – with FttA (Extended Data Figure 2; FttA is the ~73.5 kDa protein product of gene TK1428) and monitored transcription termination by quantifying release of transcripts from TECs (Figure 1). TECs stalled by nucleotide deprivation with +125 nucleotide (nt) nascent transcripts ($TECs_{+125}$) remain stably associated in the absence of FttA (Figure 1a, lanes 1-4). Addition of FttA to stalled TECs results in cleavage and release of ~100 nts of the nascent transcript (Figure 1a, lanes 5-6). However, despite repeated and exhaustive efforts to monitor FttA-mediated transcript cleavage within seconds of FttA-addition, we never observed a ~25nt 3'-transcript fragment. We were thus initially hesitant to assume that FttA-mediated transcript cleavage was coupled to *bona fide* transcription termination, as a ~25 nt transcript is sufficient to stabilize an archaeal TEC[2,23,24].

To fully validate that the cleavage and termination activity of FttA is distinct from that of a general RNase, we challenged TECs with either FttA or RNase I_f in parallel. If TECs remain intact following FttA-mediated cleavage of the nascent transcript then i) radiolabeled 3'-nascent transcripts should remain associated with TECs, ii) intact TECs should survive washes designed to remove transcripts not associated with TECs, iii) NTP addition should permit continued elongation of active TECs, allowing extension of the nascent transcript, and iv) RNAP should remain within TECs. In contrast, if FttA-mediated cleavage of the transcripts inactivates and terminates transcription, RNAP should be released to the supernatant and resumed elongation following NTP supplementation will not be possible.

Treatment of $TECs_{+125}$ with RNase I_f [10] (Figure 1c) fulfills all of the expectations of transcript processing that are not linked to transcription termination: stable $TECs_{\sim+25}$ are observed (Figure 1b, lanes 10-11), addition of unlabeled NTPs (Figure 1b, lane 12) or radiolabeled NTPs (Figure 1b, lane 13) permits all $TECs_{\sim+25}$ to resume elongation, and RNAP partitioning confirms essentially all TECs remain intact (Figure 1e, lanes 7-8). Treatment of identically prepared $TECs_{+125}$ with FttA (Figure 1b-e) are, in contrast, fully supportive of FttA-mediated termination: the bulk of FttA-treated $TECs_{+125}$ do not survive washes and FttA-activity releases ~70% of RNAP to solution (Figure 1e, lanes 5-6). FttA is thus the second archaeal-encoded factor that can mediate transcription termination.

Addition of an FttA variant [15,17,25,26] - $FttA^{H255A}$ - reduced but did not eliminate FttA-mediated termination (Figure 1a, lanes 7-8). Termination activity is thus linked to FttA-mediated RNA cleavage, rather than FttA-mediated stimulation of the intrinsic cleavage activity of RNAP[27]. FttA-mediated cleavage of the nascent RNA to yield a ~100 nt 5'-transcript is consistent with FttA stimulating RNA cleavage at the first solvent accessible phosphodiester linkage and the ~25 nt of nascent transcript protection are consistent with the results of previous digestions of intact archaeal[10] and eukaryotic TECs[28] with RNA exonucleases. In contrast to other prokaryotic transcription termination factors, FttA-mediated termination is not energy-dependent (Figure 1f).

FttA recognizes TECs through binding to nascent transcripts (Figure 2). TECs stalled on G-less cassettes, and thus with A, U and C-rich RNAs, revealed a near-linear relationship between transcript length and FttA-mediated termination (Figure 2b-d). While FttA-mediated termination is possible with only short segments of solvent accessible nascent transcript sequences – a notable discriminating feature between rho- and FttA-mediated termination – the efficiency and rate of FttA-mediated termination are modest in such instances. In contrast, TECs stalled on C-less cassettes, and thus with A, U and G-rich RNAs effectively abolish FttA-activity (Figures 2 and Extended Data Figure 3a and b). FttA-mediated termination is thus stimulated by C-rich RNAs – as is the case for bacterial rho-mediated termination – or is inhibited by transcripts that are particularly G-rich. Rho-activity can be stimulated at suboptimal *rut*-sites by NusG, and the archaeal-eukaryotic homologue of NusG, Spt5, together with its common binding partner Spt4, can likewise stimulate FttA when transcript sequences limit FttA-recognition or FttA-activity (Figure 2c-d). As such, Spt4-Spt5 temper the nucleotide requirements of FttA. FttA is a known endo- and 5'−3' exonuclease [15,29] and cleavage of nascent transcripts is stimulated by interactions with the archaeal TEC, but not RNAP alone (Extended Data Figure 3d). FttA-mediated cleavage of TEC-associated nascent transcripts is complete within ~1-2 minutes, while incubations of FttA with purified RNA under identical conditions require ~30-times longer to generate even mild cleavage patterns (Figure 3c), consistent with previous results [15,30]. FttA-mediated endonucleolytic cleavage of free RNA at CA and CC dinucleotide sequences is consistent with FttA activity on C-rich transcripts [15]. The consistently observed cleavage pattern on various substrates (reduced transcript length by ~20-30 nts) supports FttA-mediated cleavage and termination being dictated and positioned by RNAP-RNA interactions near the stalk-domain and RNA exit channel and is further enhanced by Spt4-Spt5.

Archaeal transcription units are typically separated by only short (< 100 bp) intergenic regions [1,31-33], thus for FttA-mediated termination to be effective mechanism of gene regulation *in vivo*, FttA must quickly recognize and disrupt TECs before transcription continues significantly into downstream genes or operons. To establish if FttA-mediated termination was competitive with transcription elongation, stalled $TECs_{+125}$ were permitted to resume elongation with different [NTPs] . Differential elongation rates resultant from varying [NTPs] provide a relative measure of the efficiency of FttA-mediated transcription termination in competition with transcription elongation (Figure 3). At low [NTPs], TECs elongate slowly and many TECs are still transcribing after several minutes of incubation as evidenced by a mixture of nascent transcripts between 125-225 nts (Figure 3b, lanes 3-4). At increasingly higher [NTP], elongation rates increased until TECs are elongating at rates comparable to normal elongation rates *in vivo* (Figure 3b, lanes 5-8). Addition of FttA to stalled TECs (Figure 3b, lanes 17-18) resulted in near complete termination, but as the rate of elongation increased with increasing [NTP], FttA-mediated termination decreased. Although not an obligate subcomplex of archaeal RNAP, Spt4-Spt5 engages RNAP *in vivo* early during elongation and remains associated with TECs throughout long genes[34]. The ability of Spt4-Spt5 to temper the transcript requirements for FttA-mediated termination (Figure 2) suggested that addition of Spt4-Spt5 may accelerate FttA-recognition of or action towards TECs. In support of this hypothesis, addition of Spt4-Spt5 greatly increased the termination efficiency of FttA, demonstrated by the release of transcripts > +125 nt but < +225 nt (Figure 3b, lanes 27-32). The results demonstrate that FttA is kinetically coupled to RNAP via elongation factors Spt4-Spt5, a striking analogy to the stimulation of the unrelated bacterial rho protein by NusG[35] and to the observed stimulation of Pol II termination by CPSF73/Xrn2[36]. To ensure that FttA mediates termination when combined with Spt4-Spt5 – and that termination observed in the presence of all three factors was not a new activity of Spt4-Spt5 – we employed a variant of FttA ($FttA^{H255A}$) that retains only partial activity (Extended Data Figure 4a and b).

Interactions between rho and the C-terminal KOW domain of NusG stimulate rho-mediated termination [37-39]. The NusG-KOW domain is normally engaged with the ribosome and only becomes available when transcription becomes uncoupled from translation [40]. Archaeal transcription and translation are coupled [8], and we asked whether the isolated KOW domain of Spt5 would suffice to stimulate FttA-mediated termination. Addition of the Spt5-KOW-domain ($Spt5^{\Delta NGN}$ which remains thermostable) alone does not influence the activities of FttA or RNAP *in vitro* (Extended Data Figure 4c). Spt5 is often in a heterodimeric partnership with Spt4, and this partnership is critical to kinetically couple FttA activity to RNAPs (Extended Data Figure 4d). Like the nuclear eukaryotic RNAPs, the archaeal RNAP contains a stalk domain[41]. The stalk provides binding surfaces for conserved initiation and elongation factors and the nascent transcript [1,33,42-44]. Purified stalk-less RNAP ($RNAP^{\Delta E/\Delta F}$), when combined with TBP and TFB, is competent for transcription initiation, elongation, and intrinsic termination [21], but fails to respond correctly to FttA-mediated termination (Extended Data Figure 4e). Even when continued elongation was prohibited, the termination activities of FttA were stunted by the loss of the RNAP-stalk domain (Extended Data Figure 4e, lanes 17-18). Addition of Spt4-Spt5 to TECs assembled with $RNAP^{\Delta E/\Delta F}$ does not stimulate FttA-mediated termination to rates that are competitive with continued

elongation even at low [NTP] (Extended Data Figure 4e, lanes 27-32). The results suggest roles for both the stalk domain and Spt5-Spt4 in accelerating FttA-mediated termination to permit kinetically competitive termination *in vitro*.

FttA is likely sufficiently abundant (~2,100 −/+ 500 molecules/cell; Extended Data Figure 5) to monitor global transcription (RNAP is estimated at ~3,000 molecules/cell)[19]. FttA is a metallo-beta lactamase-fold protein containing a beta-CASP domain[15] and we predicted that general inhibitors of metallo-beta-lactamase proteins[45] may impair FttA activity. 2,6-pyridine dicarboxylic acid (dipicoloinic acid, or DPA) nearly-completely inhibited FttA-mediated termination *in vitro* (Figure 4a-b). *T. kodakarensis* cultures challenged with DPA demonstrated a reduced, then complete inhibition of growth (Figure 4c). While DPA may impact several factors *in vivo*, we rationalized that monitoring RNA 3'-ends following DPA addition may reveal changes due to inactivation of FttA. Quantitative, reverse-transcription-PCR (q-RT-PCR) analyses revealed significant changes to 3'-ends at several loci *in vivo* (Figure 4d-f) following DPA addition. The fold-changes in extended 3' ends of each transcript generally increase in magnitude compared to untreated cultures both with respect to the distance from the translation stop codon and with increasing [DPA]. Altered 3'-termini stemming from exposure to a general metallo-beta-lactamase inhibitor cannot be definitively attributed to direct inhibition of FttA activity *in vivo*. Given our desire to directly demonstrate that reduced FttA activity impacts termination *in vivo*, coupled with our inability to generate *T. kodakarensis* strains encoding enzymatically-impaired FttA variants, we next reduced FttA activity by limiting FttA expression and altering steady-state $FttA^{WT}$ protein levels. To ensure that introduced and regulated expression of TK1428 did not impact TK1429 expression[46], we separated TK1428 expression from TK1429 by introduction of a new promoter and intrinsic termination sequence, then placed the TK1428 coding sequences downstream of sequences encoding an archaeal fluoride-responsive riboswitch[47], thereby generating strain IR5 (Figure 4g). Construction of IR5 was only possible when cultures were continuously provided with fluoride even though fluoride impairs growth of *T. kodakarensis*[47], supporting that very limited expression of TK1428 was not compatible with life. Steady-state FttA levels in IR5 strains grown in the absence and presence of fluoride reveals a modest ~2-fold change in FttA levels *in vivo* when fluoride is removed from cultures (Figure 4h and Supplementary Information Figure 2), yet even this modest alteration significantly and reproducibly impacts transcription termination *in vivo* (Figure 4d-f). The increased abundance of RNA with extended 3'-UTRs in strains with reduced FttA protein abundance is supportive of FttA normally directing transcription termination *in vivo*.

FttA is conserved in all archaeal genomes [13,15,48,49], including the severely reduced genomes of symbiotic Nanoarchaeota, and it was perhaps not surprising that exhaustive attempts to delete or generate variants that radically impair activity of FttA[26,50,51] in *T. kodakarensis* were unsuccessful; our failures were supported by the essentiality of FttA in other archaea[48,49]. We were able to generate a strain (termed TK1428D) encoding an His_6-affinity and HA-epitope tagged FttA(Figure 4g). Strain TK1428D growth was indistinguishable from the parental strain [52,53] and N-terminally tagged FttA was easily recovered directly from TK1428D cell lysates in large abundance (Extended Data Figure 6). Proteins co-purifying with FttA from TK1428D were identified by MuDPIT[54,55], returning only a small number of proteins (Extended Data Figure 7) that have minimal inferred

activity related to transcription and gene expression. No obvious stoichiometric FttA interaction partners were recovered, supportive of our *in vitro* demonstration that FttA alone can disrupt archaeal TECs. Affinity purification of FttA does not return RNAP subunits nor Spt4-Spt5, suggesting FttA transiently encounters and disrupts TECs rather than forming stable interactions with TEC components or Spt4-Spt5.

The essentiality of FttA in *T. kodakarensis* and other archaea[48,49], the complete conservation of FttA in Archaea[15,18,56], the demonstrated *in vitro* ability of FttA-mediated termination to compete with productive elongation (Figure 3) and the demonstrated changes to RNA 3'-ends in strains wherein FttA activity is reduced by two independent mechanisms (Figure 4) suggests that FttA is likely responsible for 3'-end formation of transcripts that are not directed by intrinsic termination, and further that FttA-mediated termination is likely responsible for polarity in archaeal cells [32]. By establishing the requirements for FttA-mediated transcription termination (Figures 1-4, Extended Data Figure 3 and Extended Data Figure 4) we complete the archaeal transcription cycle and describe an additional mechanism of 3' end formation (Extended Data Figure 8). It is important to note that the described activities of FttA suggest that the steady-state 3'-termini of *in vivo* transcripts terminated by FttA do not reflect that actual position of termination of the archaeal RNAP. Thus, consensus termination sequences derived from next-generation sequencing and Term-seq data[1,57] should be re-evaluated given that FttA-terminated transcripts are likely to be lacking minimally ~20-30 nts from the 3'-terminus; additional RNA processing events are likely to further complicate attempts to map the 3'-termini of transcripts that reflect the true position of TEC dissociation. Even transcripts derived from loci encoding putative intrinsic termination sequences should be reevaluated, as FttA-activity may influence the efficiency of intrinsic termination or serve as a backup mechanism of transcription termination for genes/operons with less-efficient intrinsic termination signals.

The requirements for FttA-mediated termination suggest that long 5' UTRs observed for some archaeal transcripts may serve as points of regulation for premature termination upstream of the coding sequences [46,58]. How transcription of stable RNAs, including rRNAs is protected from FttA-mediated termination will be of interest to determine. Exclusion of Spt4-Spt5 from TECs transcribing stable RNAs, or structures within the nascent transcript may suffice to hinder FttA-loading or FttA-mediated termination of archaeal TECs; a delayed mechanism of Spt5 recruitment to rRNA and CRIPSR loci has been identified [59]. It is interesting that full-length FttA homologues are found in the genomes of several bacterial species, suggesting that FttA may function as a termination factor in multiple domains (Supplementary Information Figure 1). It will be of immediate interest to determine if the bacterial FttA proteins can direct transcription termination, and if they can, whether they cooperate with or can substitute for rho. It will be similarly interesting to determine if FttA activity can disrupt eukaryotic TECs formed with Pol I, II and III, given the structure of FttA is nearly identical to the CPSF73 subunit of the eukaryotic CPSF complex (Extended Data Figure 2) [15,17,25,60,61]. The combined activities of CPSF and Xrn2 are necessary for normal termination patterns in Eukarya [62,63]. FttA retains all the necessary activities within a single protein: FttA can bind TECs, mediate cleavage and release of the nascent transcript and use 5'−3' exonuclease activities to degrade the 3'-transcript. We propose that the eukaryotic CPSF complex [64], which minimally contains 4 homologous but non-identical subunits, arose

from archaeal FttA. The ability of the CSPF complex to directly terminate transcription was likely lost during the specialization and partnership with factors that direct RNA 3'-maturation in Eukarya.

Methods

T. kodakarensis culturing conditions.

T. kodakarensis strain TS559 and derivatives of such were grown at 85°C under anaerobic conditions as previously described[53]. DPA (Sigma) was added at neutral pH to either 12.5 or 25 mM as shown in Figure 4. NaF was added to 4 mM when necessary.

Protein purifications.

Archaeal RNA polymerases (WT and ΔE/F variant) containing His_6-HA-epitope-tagged-RpoL subunits, TBP, and TFB were purified as previously described [19,21]. *T. kodakarensis* Spt5 and His_6-Spt4 were purified as previously described [20]. $Spt5^{\Delta NGN}$ was purified as was full-length Spt5. WT and an H255A variant of FttA were purified from Rosetta2 *E. coli* cells carrying pQE-80L (Qiagen) expression vectors carrying the wildtype or variant TK1428 coding sequence (Extended Data Figure 2). Cells were grown in LB medium at 37°C with shaking (~220 rpm) with 30 μg/ml chloramphenicol and 100 μg/ml ampicillin to an optical density at 600 nm of 0.5 before expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside. Cultures were grown for an additional 3 h at 37°C with shaking before biomass was harvested via centrifugation (~8,000 x *g*, 20 min, 4°C), resuspended and lysed via sonication (3 ml/g of biomass) in 20 mM Tris-HCl pH 8.0, 5 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 100 mM NaCl. Cellular lysates were clarified by centrifugation (~20,000 x *g*, 20 min, 4°C), heated to 85°C for 30 min to denature most host proteins, and clarified again by centrifugation (~20,000 x *g*, 20 min, 4°C). Heat-treated clarified cell lysates were resolved through a 5 ml HiTrap-heparin column (GE Healthcare) with a linear gradient from 0.1 – 1.0 M NaCl dissolved in 20 mM Tris-HCl pH 8.0, 5 mM 2-mercaptoethanol, 10 mM $MgCl_2$. Fractions containing > 95% pure FttA were identified by SDS-PAGE, pooled, and dialyzed into 25 mM Tris-HCl pH 8.0, 100 mM KCl, 10 mM 2-mercaptoethanol, 50% glycerol before storage at −80°C. All protein concentrations were quantified using a Bradford Assay [65].

DNA templates.

Double-stranded DNA templates used in all transcription reactions were PCR amplified from plasmids and gel purified as previously described [19,20,24]. All transcription templates contain a non-template 5'-strand biotin-TEG moiety to provide attachment to streptavidin-coated paramagnetic beads (Promega).

In vitro transcription assays.

Assembly of preinitiation complexes (PICs) and elongation via NTP deprivation was carried out as described previously [19,20,24]. To obtain stalled TECs on G-less cassette templates, PICs were assembled using 10 nM template, 20 nM RNAP, 40 nM TBP, 40 nM TFB in a 20 μl total volume of transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) with 75 μM ApC for 3 min at 85°C before addition of 200 μM ATP,

200 μM CTP, 10 μM UTP and 10 μCi [α−^{32}P]-UTP for 3 additional min at 85°C, then chilled to 4°C. To obtain stalled TECs on C-less cassette templates, reactions were identical to those above, with the substitution of 200 μM GTP for 200 μM CTP. RNAP bound templates were captured with HisPur™ Ni-NTA magnetic particles (Thermo Fisher Scientific) and washed three times with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl.

For Figure 1a, washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with 10 μM each of ATP, CTP, and UTP before addition of 1 μM FttA or FttA^{H255A} for 5 min at 85°C. Reactions were chilled to 4°C followed by separation of pellet and supernatant fractions by addition of streptavidin coated paramagnetic particles (Promega). Similar results were observed in 4 independent experiments. For Figure 1b, washed TECs (lane 1) were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT with 10 μM each of ATP, CTP, and UTP and were incubated at 85°C for 7 min (lane 2) before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 3). Washed TECs were incubated at 85°C for 1 minute before addition of 100 μM NTPs (lane 4) or 100 μM ATP, CTP and GTP, 10μM UTP containing 1μC ^{32}P-α-UTP (lane 5) and continued incubation at 85°C for 3 minutes. Washed TECs were exposed to 1μM FttA (lane 6) at 85°C for 7 minutes before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 7). FttA-treated, washed TECs were incubated at 85°C for 1 minute before addition of 100 μM NTPs (lane 8) or 100 μM ATP, CTP and GTP, 10μM UTP containing 1μC ^{32}P-α-UTP (lane 9) and continued incubation at 85°C for 3 minutes. Washed TECs were exposed to 50U RNase I_f (lane 10) at 37°C for 7 minutes before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 11). RNase I_f-treated, washed TECs were incubated at 85°C for 1 minute before addition of 100 μM NTPs (lane 12) or 100 μM ATP, CTP and GTP, 10μM UTP containing 1μC ^{32}P-α-UTP (lane 13) and continued incubation at 85°C for 3 minutes. Similar results were observed in 4 independent experiments. For Figure 1, panel f, washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with or without 10 μM each of ATP, CTP, and UTP before addition of 1 μM FttA at 85°C. Reaction aliquots were removed after 1, 2, or 5 minutes, chilled to 4°C, then pellet and supernatant fractions were separated by addition of streptavidin coated paramagnetic particles (Promega). Similar results were observed in 4 independent experiments.

For Figures 2, panels b and c, and Extended Data Figure 3, washed TECs were assembled as above on G-less or C-less cassettes of various lengths. Washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with or without 1 μM FttA at 85°C for 3 minutes. Reactions were chilled to 4°C, then pellet and supernatant fractions were separated by addition of streptavidin coated paramagnetic particles (Promega). For Figure 2b-c and Extended Data Figure 3 similar results were observed in 3 independent experiments.

For Figures 3 and Extended Data Figure 4, washed TECs were assembled with WT or ΔE/F RNAP as above on a +125 G-less cassette template. Continued elongation was permitted by the addition of 0, 1, 10 or 100 μM NTPs in the presence or absence of combinations of 6 μM Spt4, 6 μM Spt5, 6 μM $Spt5^{\Delta NGN}$, 1 μM FttA or $FttA^{H255A}$. After 5 min at 85°C, reactions were chilled to 4°C followed by separation of pellet and supernatant fractions by addition of streptavidin coated paramagnetic particles (Promega). For Figure 3c, similar results were observed in 3 independent experiments. Extended Data Figure 4 was performed once.

For Figure 4b, stalled TECs on a G-less cassette template assembled using 10 nM template, 20 nM RNAP, 40 nM TBP, 40 nM TFB in a 20 μl total volume of transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) with 75 μM ApC for 3 min at 85°C before addition of 200 μM ATP, 200 μM CTP, 10 μM UTP and 10 μCi [α-^{32}P]-UTP and incubation for 3 additional min at 85°C, then chilled to 4°C. RNAP bound templates were captured with HisPur™ Ni-NTA magnetic particles (Thermo Fisher Scientific) and washed three times with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl. Washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with 10 μM each of ATP, CTP, and UTP before addition of reaction buffer (10 mM Tris-HCl pH 8.0, 120 mM KCl, 8 mM DTT, and 1.25 mM $MgCl_2$) +/− 25 mM DPA and/or +/− 1 μM FttA at 85°C. Reaction aliquots were incubated for 3 minutes, and then chased with 250 μM ATP, CTP, UTP, and GTP for 2 minutes to allow for elongation to +225. Reactions were chilled to 4°C, then pellet and supernatant fractions were separated by addition of streptavidin coated paramagnetic particles (Promega). Similar results were observed in 5 independent experiments.

Radiolabeled transcripts from Figures 1-4, Extended Data Figure 3 and Extended Data Figure 4 were recovered by addition of 5 volumes STOP buffer (600 mM Tris-HCl pH 8.0, 30 mM EDTA) containing 7 μg of tRNA (total), equal volume phenol/chloroform/isoamyl alcohol (25:24:1, V/V) extractions, and precipitations of the aqueous phase with 2.6 volumes 100% ethanol. Precipitated transcripts were resuspended in 95% formamide, 1X TBE, heated to 99°C for 2 minutes, rapidly chilled on ice, loaded and resolved in 10-20% polyacrylamide, 8M urea, 1X TBE denaturing gels. Radiolabeled RNA was detected using phosphorimaging (GE Healthcare). Gel images were analyzed using GE Imagequant 5.2 software.

To generate +125 nt [α-^{32}P]-UTP labeled transcripts used in Extended Data Figure 3, transcription reactions were assembled and terminated as above, with the substitution of 10 μg of glycogen for 7 μg of tRNA during reaction clean up. Radiolabeled transcripts were incubated at 85°C with or without 1 μM FttA or at 37°C with 5U RNaseA (Thermo Fisher Scientific) in *T. kodakarensis* transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) to monitor cleavage. When present, *T. kodakarensis* RNAP was added to 40 nM final.

Western blot analysis of RNAP release to solution (Figure 1, panel e).

Anti-HA antibodies (BioLegend 901513) were employed as previously described[10,19,53] to quantify RpoL-levels in P and S fractions.

Western blot analysis of FttA protein levels (Figure 4, panel h, and Extended Data Figure 5).

Purified, recombinant full-length FttA was used as an antigen to prepare polyclonal antibodies in mice (Cocalico Biologicals). Known quantities of purified FttA were resolved as comparative quantification standards in adjacent lanes to clarified cell lysates derived from known quantities of cells. Proteins were separated via SDS-PAGE, transferred to PVDF membranes, and probed with primary anti-FttA antibodies. Blots were developed by addition of an IgG-AP conjugated anti-mouse secondary antibody allowed for detection by NBT/BCIP (Roche). A linear regression FttA signal intensity to FttA amount in ng was generated.

Construction of strains TK1428D and IR5.

T. kodakarensis strains used here were constructed from the parental strains TS559[52] as previously described[53].

Purification of FttA directly from lysates of strain TK1428D and MuDPIT analysis.

Purification procedures and MuDPIT analysis was performed largely as previously described[10,54,55]. Briefly, 2 L early-exponential phase cultures ($OD_{600\ nm}$ = ~0.3) of *T. kodakarensis* strains TK1428D and TS559 grown in ASW-YT-S° medium at 85°C were rapidly chilled to 4°C and harvested by centrifugation (20,000 x *g*). All subsequent procedures were carried out at 4°C and completed as rapidly as possible to retain native *in vivo* protein-protein interactions. Cells from each strain were individually resuspended in 3 ml of 25 mM Tris-HCl pH 8, 500 mM NaCl, 10% glycerol per gram of wet biomass and lysed by repetitive sonication. The resulting lysates were clarified via centrifugation (20,000 x *g*) before passage through 5-ml HiTRAP chelating columns (GE Healthcare) pre-charged with $NiSO_4$ and equilibrated in 25 mM Tris-HCl pH 8, 500 mM NaCl, 10% glycerol. The lysate components that did not bind the columns (e.g. flowthrough) were discarded and the columns were washed with ~20 column volumes 25 mM Tris-HCl pH 8, 500 mM NaCl, 10% glycerol until no additional proteins eluted. Bound proteins were eluted using a linear gradient from (initial) 25 mM Tris-HCl pH 8, 500 mM NaCl, 10% glycerol to (final) 25 mM Tris-HCl pH 8, 100 mM NaCl, 150 mM imidazole, 10% glycerol. Fractions that contained the tagged TK1428 protein were identified by Western blotting (Extended Data Figure 6), pooled, and were twice dialyzed in 3 Kda molecular weight cut-off dialysis tubing against 1 L 25 mM Tris-HCl pH 8, 500 mM NaCl, 0.5 mM EDTA, 2 mM dithiothreitol. FttA was not identified by Western blots within fractions resultant from processing TS559 biomass, but identical fractions were collected and processed to identify native *T. kodakarensis* present in such fractions that spontaneously bind the chelating resin. Protein concentrations were determined by Bradford assays[65], and thirty-microgram aliquots of the proteins present in solution were precipitated by adding trichloroacetic acid (TCA; 15% final concentration). The TCA-precipitated proteins were identified by multidimensional protein identification technology at the Ohio State University mass spectrometry facility (https://www.ccic.osu.edu/MSP) using the MASCOT search engine. We required a minimum of two unique peptide fragments to positively identify a protein. TS559 control samples identified several proteins that bound and eluted from the Ni^{2+}-charged matrix in the absence of a

His_6-tagged protein. All of the proteins identified in the experimental samples that had MASCOT scores of >100 and were not also present in the control samples are shown in Extended Data Figure 7.

TRIzol-based RNA purifications from *T. kodakarensis* cultures and q-RT-PCR.

RNA extractions were performed essentially as previously described[46] from strain TS559 prior to, or 1 hour after DPA addition. RNA extractions from IR5 and TS559, in the absence and presence of 4 mM NaF were performed as previously described[46]. qRT-PCR reactions were performed as previously described[66], except that 500 ng of total RNA was used during cDNA synthesis.

Sequences of DNA templates employed in transcription reactions.

For all DNA templates, the BRE and TATA sequences are shown in capital text and italicized, transcription start sites are in capital bold text, transcribed regions are underlined, and positions used to stall TECs by NTP deprivation are shown in bold. Critical parameters of each template (template name, G-less or C-less cassette length and the length of full-length run-off transcripts) are highlighted above the text of each DNA templates

TJS0 – 125nt G-less stalled transcript, 175nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatctctcttctcaaata
ttctacttattactcaacaccatctaccttccactccatcctaccttccacatctctcttctcaccctccctctccaccaactacttccacttca
ccatcaaccatgggtcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgg

TJS1 – 125nt G-less stalled transcript, 225nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatctctcttctcaaata
ttctacttattactcaacaccatctaccttccactccatcctaccttccacatctctcttctcaccctccctctccaccaactacttccacttca
ccatcaaccatgggtcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacc
ctttagtgagggttaattgcgg

TJS2 – 87nt G-less stalled transcript, 187nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatctctcttctcaaata
ttctacttattactcaacaccatctaccttccactccatcctaccttccacatctctcttctcaccctgggtcctgttccgttggaccgtggttt
ccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacccctttagtgagggttaattgcgg

TJS3 – 69nt G-less stalled transcript, 169nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatctctcttctcaaata
ttctacttattactcaacaccatctaccttccactccatcctaccttccagggtcctgttccgttggaccgtggtttccaggcggtgcgtct
gctgtcgggtccctctggtgactggctagatggacccctttagtgagggttaattgcgg

TJS4 – 54nt G-less stalled transcript, 154nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatctctcttctcaaata
ttctacttattactcaacaccatctaccttccactgggtcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctc
tggtgactggctagatggacccctttagtgagggttaattgcgg

TJS5 – 125nt C-less stalled transcript, 180nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatgtgtgttgtgaaat

attgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagtgttaatgtggttgtagataaagttaatgtatat tggtaatatgagttccctcctgttcgtcgccgttggaccgtggttccctttagtgagggttaattgcgg

TJS6 – 98nt C-less stalled transcript, 153nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatgtgtgttgtgaaat attgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagtgttaatgtggttgtagataaccctcctgttcgt cgccgttggaccgtggttccctttagtgagggttaattgcgg

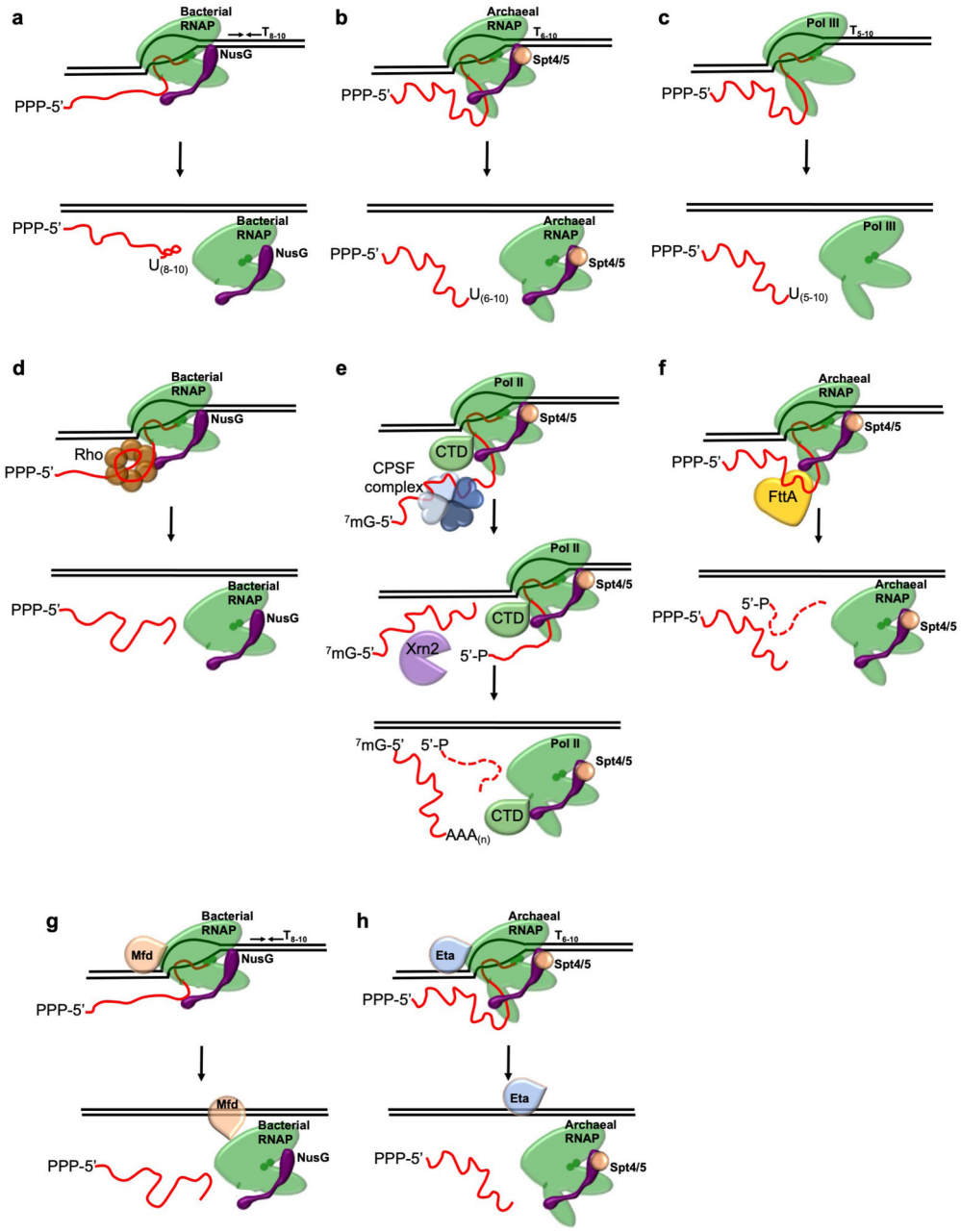
TJS7 – 78nt C-less stalled transcript, 133nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatgtgtgttgtgaaat attgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagtccctcctgttcgtcgccgttggaccgtggtt ccctttagtgagggttaattgcgg

TJS8 – 58nt C-less stalled transcript, 113nt run-off transcript
cgcgcgtaatacgactcactataggg*GCGA*tata*TTTATATA*gggatatagtaatagataatatcAcatgtgtgttgtgaaat attgtagttggatgtaggttggagtggagtgtgtggtgtgccctcctgttcgtcgccgttggaccgtggttccctttagtgagggttaatt gcgg

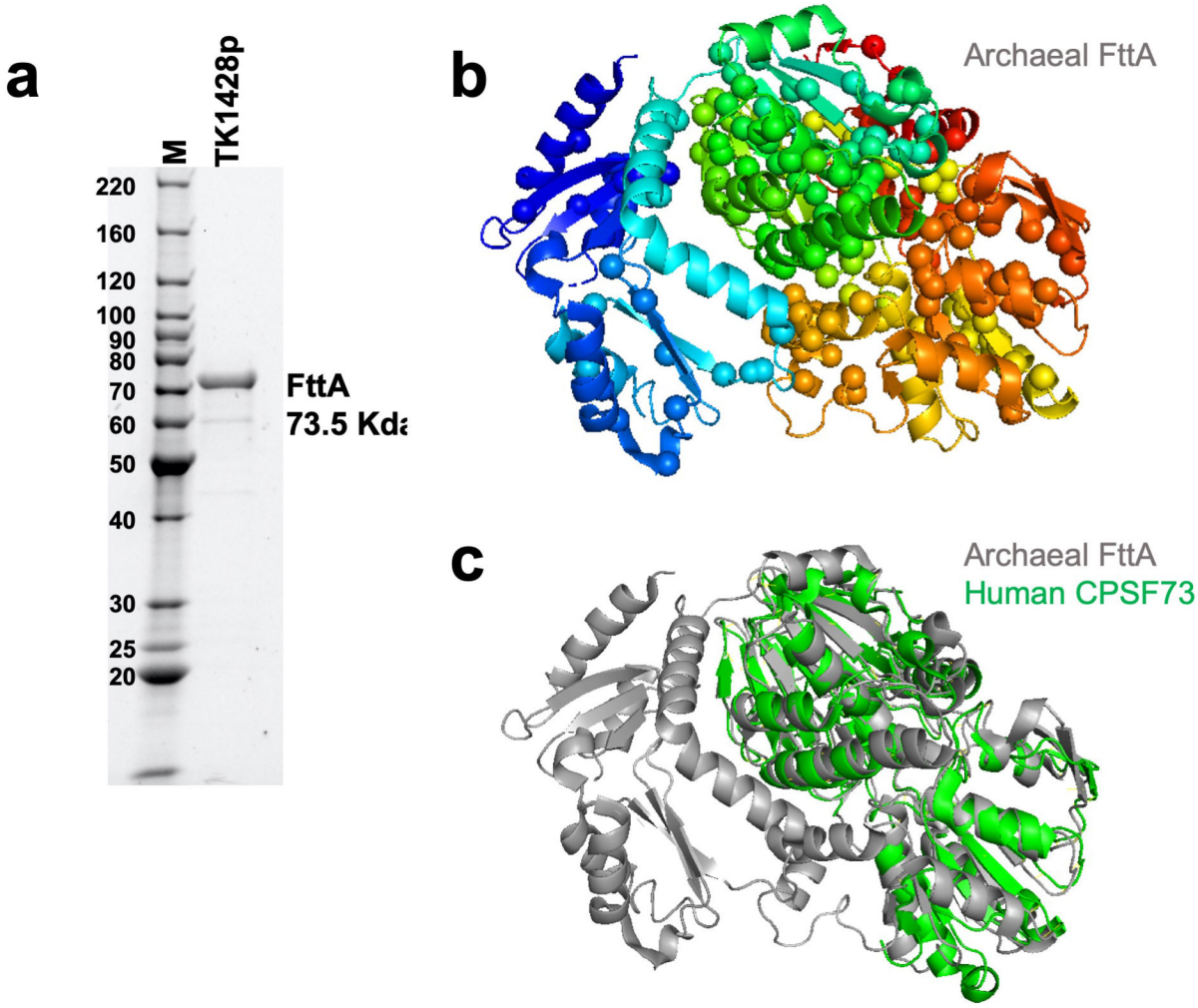
Data availability.

The data that support the findings of this study are available from the authors upon reasonable request. Source data for Figs. 1e, and 4h, as well as Extended Data Figs. 2a, 3d and 5 are included in this article and its Supplemental Information Files.

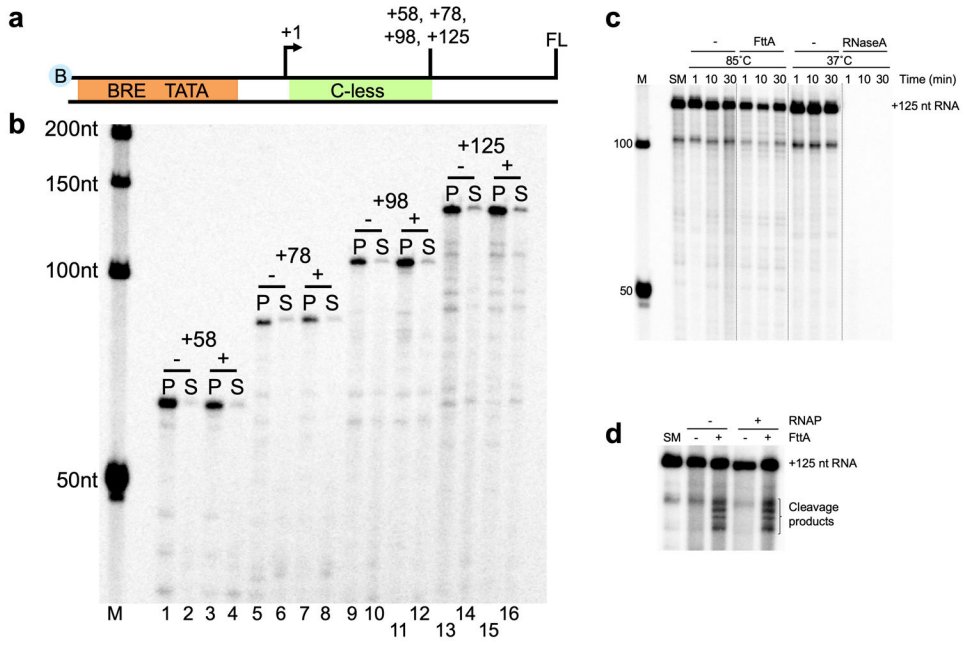
Extended Data



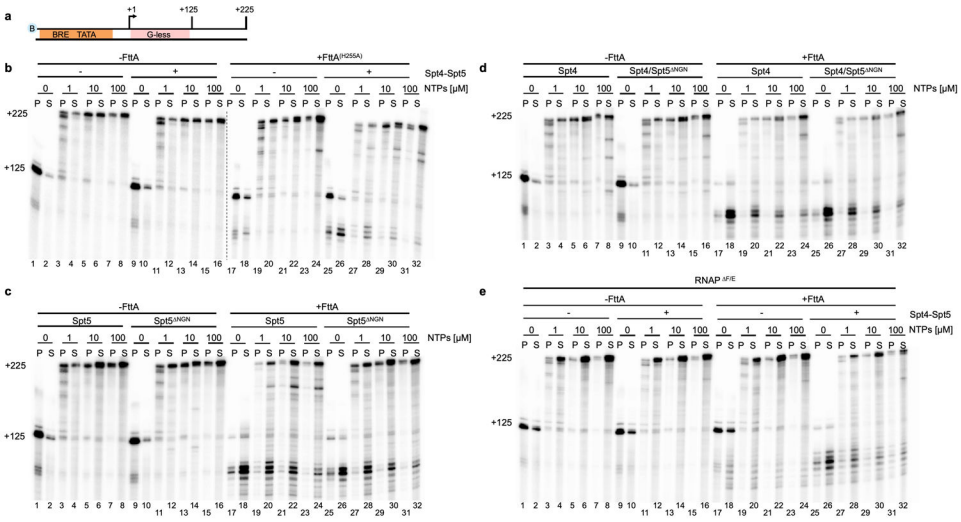
Extended Data Figure 1.



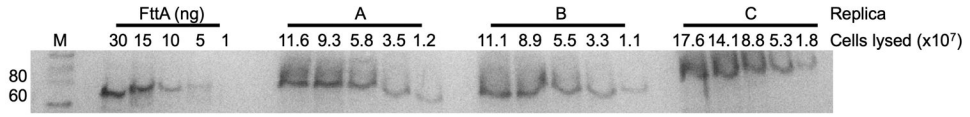
Extended Data Figure 2.



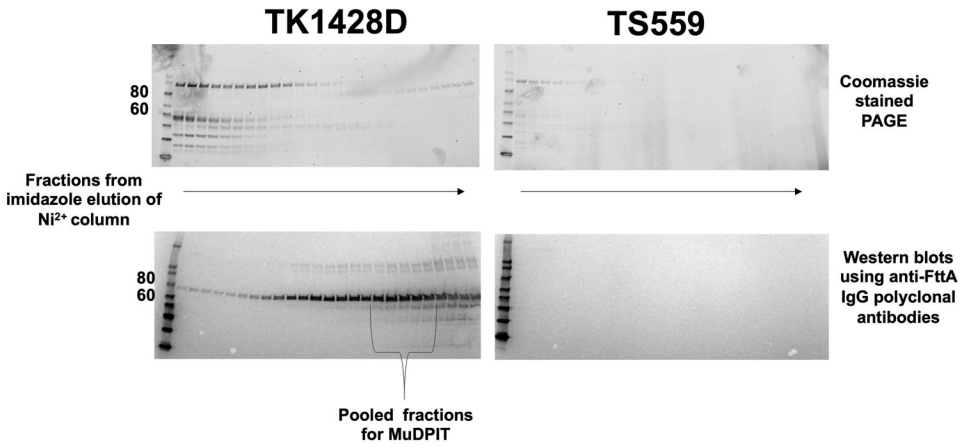
Extended Data Figure 3.



Extended Data Figure 4.



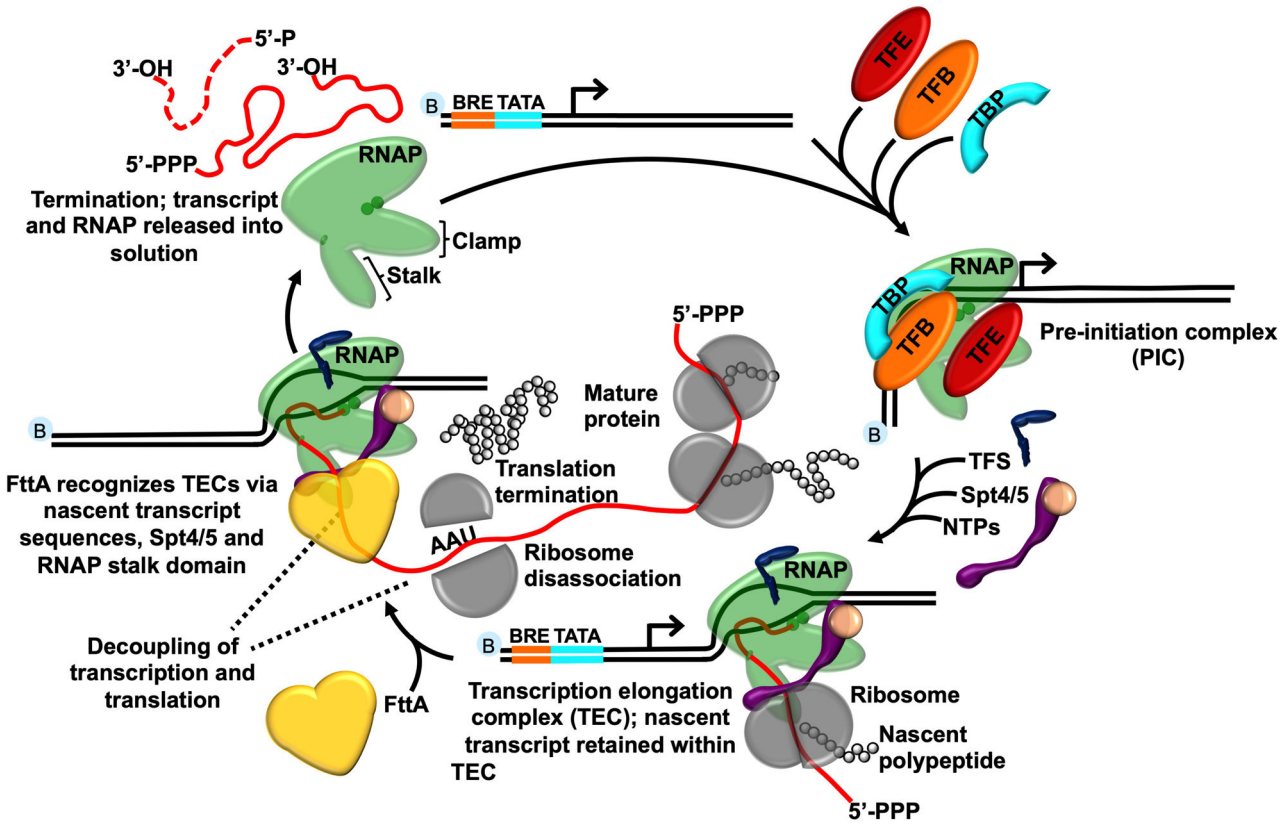
Extended Data Figure 5.



Extended Data Figure 6.

Gene	Annotation	Mascot Score	Number of Unique Peptides
TK 1428	Cleavage and Polyadenylation specificity factor homologue	6388	47
TK 2215	tRNA splicing endonuclease	633	12
TK 0011	Uncharacterized protein	585	7
TK 1557	Predicted dehydrogenase	537	18
TK 1165	Predicted AP endonuclease	468	18
TK 2250	Serine/Threonine protein kinase	421	17
TK 0976	Putative snRNP Sm-like protein	351	8
TK 1509	Probable tRNA pseudouridine synthase	202	13
TK 0528	Serine hydroxyethyltransferase	168	13
TK 0211	Amidophosphoribosyltransferase	147	9
TK 1305	Probable translation initiation factor IF-2	133	11

Extended Data Figure 7.



Extended Data Figure 8.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the lab for critical review of the manuscript. This work was supported by NIH grant GM100329 (to T.J. Santangelo).

References:

1. Maier L-K & Marchfelder A It's all about the T: transcription termination in archaea. Biochem. Soc. Trans 47, 461–468 (2019). [PubMed: 30783016]
2. Santangelo TJ & Reeve JN Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J. Mol. Biol 355, 196–210 (2006). [PubMed: 16305799]
3. Sela I, Wolf YI & Koonin EV Theory of prokaryotic genome evolution. Proc. Natl. Acad. Sci. U. S. A 113, 11399–11407 (2016). [PubMed: 27702904]
4. Helmrich A, Ballarino M, Nudler E & Tora L Transcription-replication encounters, consequences and genomic instability. Nat. Struct. Mol. Biol 20, 412–418 (2013). [PubMed: 23552296]
5. Washburn RS & Gottesman ME Transcription termination maintains chromosome integrity. Proc. Natl. Acad. Sci 108, 792–797 (2011). [PubMed: 21183718]
6. Shin J-H, Santangelo TJ, Xie Y, Reeve JN & Kelman Z Archaeal minichromosome maintenance (MCM) helicase can unwind DNA bound by archaeal histones and transcription factors. J. Biol. Chem 282, 4908–15 (2007). [PubMed: 17158792]

7. Miller OL, Hamkalo BA & Thomas CA Visualization of Bacterial Genes in Action. Science (80-.). 169, 392–395 (1970).
8. French SL, Santangelo TJ, Beyer AL & Reeve JN Transcription and translation are coupled in Archaea. Mol. Biol. Evol 24, 893–5 (2007). [PubMed: 17237472]
9. Ray-Soni A, Bellecourt MJ & Landick R Mechanisms of Bacterial Transcription Termination: All Good Things Must End. Annu. Rev. Biochem 85, 319–47 (2016). [PubMed: 27023849]
10. Walker JE, Luyties O & Santangelo TJ Factor-dependent archaeal transcription termination. Proc. Natl. Acad. Sci 114, E6767–E6773 (2017). [PubMed: 28760969]
11. Santangelo TJ et al. Polarity in archaeal operon transcription in Thermococcus kodakaraensis. J. Bacteriol 190, 2244–8 (2008). [PubMed: 18192385]
12. D'Heygere F, Rabhi M & Boudvillain M Phyletic distribution and conservation of the bacterial transcription termination factor Rho. Microbiology 159, 1423–1436 (2013). [PubMed: 23704790]
13. Makarova K, Wolf Y & Koonin E Archaeal Clusters of Orthologous Genes (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and Methanobacteriales. Life 5, 818–840 (2015). [PubMed: 25764277]
14. Wolf YI, Makarova KS, Yutin N & Koonin EV Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. Biol. Direct 7, 46 (2012). [PubMed: 23241446]
15. Phung DK et al. Archaeal β-CASP ribonucleases of the aCPSF1 family are orthologs of the eukaryal CPSF-73 factor. Nucleic Acids Res. 41, 1091–103 (2013). [PubMed: 23222134]
16. Mandel CR et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. Nature 444, 953–6 (2006). [PubMed: 17128255]
17. Nishida Y et al. Crystal structure of an archaeal cleavage and polyadenylation specificity factor subunit from Pyrococcus horikoshii. Proteins 78, 2395–8 (2010). [PubMed: 20544974]
18. Galperin MY, Kristensen DM, Makarova KS, Wolf YI & Koonin EV Microbial genome analysis: the COG approach. Brief. Bioinform (2017). doi:10.1093/bib/bbx117
19. Santangelo TJ, Cubonová L, James CL & Reeve JN TFB1 or TFB2 is sufficient for Thermococcus kodakaraensis viability and for basal transcription in vitro. J. Mol. Biol 367, 344–57 (2007). [PubMed: 17275836]
20. Sanders TJ et al. TFS and Spt4/5 accelerate transcription through archaeal histone-based chromatin. Mol. Microbiol doi:10.1111/mmi.14191
21. Hirata A et al. Archaeal RNA polymerase subunits E and F are not required for transcription in vitro, but a Thermococcus kodakarensis mutant lacking subunit F is temperature-sensitive. Mol. Microbiol 70, 623–33 (2008). [PubMed: 18786148]
22. Santangelo TJ & Reeve JN Deletion of switch 3 results in an archaeal RNA polymerase that is defective in transcript elongation. J. Biol. Chem 285, 23908–15 (2010). [PubMed: 20511223]
23. Gehring AM & Santangelo TJ Archaeal RNA polymerase arrests transcription at DNA lesions. Transcription e1324941 (2017). doi:10.1080/21541264.2017.1324941
24. Gehring AM & Santangelo TJ Manipulating Archaeal Systems to Permit Analyses of Transcription Elongation-Termination Decisions In Vitro. in Methods in molecular biology (Clifton, N.J.) 1276, 263–279 (2015).
25. Mir-Montazeri B, Ammelburg M, Forouzan D, Lupas AN & Hartmann MD Crystal structure of a dimeric archaeal Cleavage and Polyadenylation Specificity Factor. J. Struct. Biol 173, 191–195 (2011). [PubMed: 20851187]
26. Kolev NG, Yario TA, Benson E & Steitz JA Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. EMBO Rep. 9, 1013–8 (2008). [PubMed: 18688255]
27. Orlova M, Newlands J, Das A, Goldfarb A & Borukhov S Intrinsic transcript cleavage activity of RNA polymerase. Proc. Natl. Acad. Sci. U. S. A 92, 4596–600 (1995). [PubMed: 7538676]
28. Dengl S & Cramer P *Torpedo* Nuclease Rat1 Is Insufficient to Terminate RNA Polymerase II *in Vitro*. J. Biol. Chem 284, 21270–21279 (2009). [PubMed: 19535338]

29. Phung DK & Clouet-d'Orval B Tips and Tricks to Probe the RNA-Degrading Activities of Hyperthermophilic Archaeal β-CASP Ribonucleases. 453–466 (2015). doi:10.1007/978-1-4939-2214-7_26
30. Silva APG et al. Structure and activity of a novel archaeal β-CASP protein with N-terminal KH domains. Structure 19, 622–32 (2011). [PubMed: 21565697]
31. Ray WC & Daniels CJ PACRAT: a database and analysis system for archaeal and bacterial intergenic sequence features. Nucleic Acids Res. 31, 109–13 (2003). [PubMed: 12519960]
32. Santangelo TJ et al. Polarity in archaeal operon transcription in Thermococcus kodakaraensis. J. Bacteriol 190, 2244–8 (2008). [PubMed: 18192385]
33. Gehring AM, Walker JE & Santangelo TJ Transcription Regulation in Archaea. J. Bacteriol 198, 1906–1917 (2016). [PubMed: 27137495]
34. Smollett K, Blombach F, Reichelt R, Thomm M & Werner F A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. Nat. Microbiol 2, 17021 (2017). [PubMed: 28248297]
35. Cardinale CJ et al. Termination Factor Rho and Its Cofactors NusA and NusG Silence Foreign DNA in E. coli. Science (80-.). 320, 935–938 (2008).
36. Cortazar MA et al. Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. Mol. Cell (2019). doi:10.1016/j.molcel.2019.09.031
37. Lawson MR & Berger JM Tuning the sequence specificity of a transcription terminator. Curr. Genet (2019). doi:10.1007/s00294-019-00939-1
38. Lawson MR et al. Mechanism for the Regulated Control of Bacterial Transcription Termination by a Universal Adaptor Protein. Mol. Cell 71, 911–922.e4 (2018). [PubMed: 30122535]
39. Mitra P, Ghosh G, Hafeezunnisa M & Sen R Rho Protein: Roles and Mechanisms. Annu. Rev. Microbiol 71, 687–709 (2017). [PubMed: 28731845]
40. Burmann BM et al. A NusE:NusG Complex Links Transcription and Translation. Science (80-.). 328, 501–504 (2010).
41. Werner F & Grohmann D Evolution of multisubunit RNA polymerases in the three domains of life. Nat. Rev. Microbiol 9, 85–98 (2011). [PubMed: 21233849]
42. Nagy J et al. Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. Nat. Commun 6, 6161 (2015). [PubMed: 25635909]
43. Plaschka C et al. Architecture of the RNA polymerase II–Mediator core initiation complex. Nature 518, 376–380 (2015). [PubMed: 25652824]
44. Walker JE & Santangelo TJ Analyses of in vivo interactions between transcription factors and the archaeal RNA polymerase. Methods 86, 73–79 (2015). [PubMed: 26028597]
45. Horsfall LE et al. Competitive Inhibitors of the CphA Metallo- -Lactamase from Aeromonas hydrophila. Antimicrob. Agents Chemother 51, 2136–2142 (2007). [PubMed: 17307979]
46. Jäger D, Förstner KU, Sharma CM, Santangelo TJ & Reeve JN Primary transcriptome map of the hyperthermophilic archaeon Thermococcus kodakarensis. BMC Genomics 15, 684 (2014). [PubMed: 25127548]
47. Speed MC, Burkhart BW, Picking JW & Santangelo TJ An Archaeal Fluoride-Responsive Riboswitch Provides an Inducible Expression System for Hyperthermophiles. Appl. Environ. Microbiol 84, e02306–17 (2018). [PubMed: 29352088]
48. Sarmiento F, Mrázek J & Whitman WB Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon Methanococcus maripaludis. Proc. Natl. Acad. Sci. U. S. A 110, 4726–31 (2013). [PubMed: 23487778]
49. Zhang C, Phillips APR, Wipfler RL, Olsen GJ & Whitaker RJ The essential genome of the crenarchaeal model Sulfolobus islandicus. Nat. Commun 9, 4908 (2018). [PubMed: 30464174]
50. Garas M, Dichtl B & Keller W The role of the putative 3' end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. RNA 14, 2671–84 (2008). [PubMed: 18971324]
51. Ryan K, Calvo O & Manley JL Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. RNA 10, 565–73 (2004). [PubMed: 15037765]

52. Santangelo TJ, Cubonová L & Reeve JN Thermococcus kodakarensis genetics: TK1827-encoded beta-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology. Appl. Environ. Microbiol 76, 1044–52 (2010). [PubMed: 20023088]
53. Gehring A, Sanders T & Santangelo TJ Markerless Gene Editing in the Hyperthermophilic Archaeon Thermococcus kodakarensis. BIO-PROTOCOL 7, (2017).
54. Li Z, Santangelo TJ, Cuboňová L, Reeve JN & Kelman Z Affinity purification of an archaeal DNA replication protein network. MBio 1, e00221-10–e00221-19 (2010). [PubMed: 20978540]
55. Burkhart BW, Febvre HP & Santangelo TJ Distinct Physiological Roles of the Three Ferredoxins Encoded in the Hyperthermophilic Archaeon *Thermococcus kodakarensis*. MBio 10, (2019).
56. Wolf YI, Makarova KS, Yutin N & Koonin EV Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. Biol. Direct 7, 46 (2012). [PubMed: 23241446]
57. Dar D, Prasse D, Schmitz RA & Sorek R Widespread formation of alternative 3' UTR isoforms via transcription termination in archaea. Nat. Microbiol 1, 16143 (2016). [PubMed: 27670118]
58. Cohen O et al. Comparative transcriptomics across the prokaryotic tree of life. Nucleic Acids Res. 44, W46–W53 (2016). [PubMed: 27154273]
59. Smollett K, Blombach F, Reichelt R, Thomm M & Werner F A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. Nat. Microbiol 2, 17021 (2017). [PubMed: 28248297]
60. Sun Y et al. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. Proc. Natl. Acad. Sci. U. S. A 115, E1419–E1428 (2018). [PubMed: 29208711]
61. Clerici M, Faini M, Muckenfuss LM, Aebersold R & Jinek M Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. Nat. Struct. Mol. Biol 25, 135–138 (2018). [PubMed: 29358758]
62. Eaton JD et al. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. Genes Dev. 32, 127–139 (2018). [PubMed: 29432121]
63. Proudfoot NJ Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science 352, aad9926 (2016). [PubMed: 27284201]
64. Hill CH et al. Activation of the Endonuclease that Defines mRNA 3′ Ends Requires Incorporation into an 8-Subunit Core Cleavage and Polyadenylation Factor Complex. Mol. Cell (2019). doi:10.1016/j.molcel.2018.12.023
65. Bradford MM A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem 72, 248–54 (1976). [PubMed: 942051]
66. Mattiroli F et al. Structure of histone-based chromatin in Archaea. Science (80-.). 357, 609–612 (2017).
67. Roberts JW Mechanisms of Bacterial Transcription Termination. J. Mol. Biol (2019). doi:10.1016/j.jmb.2019.04.003
68. Mishra S & Maraia RJ RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3′ end dynamics during transcription termination. Nucleic Acids Res. 47, 310–327 (2019). [PubMed: 30407541]
69. Gehring AM, Walker JE & Santangelo TJ Transcription Regulation in Archaea. J. Bacteriol 198, 1906–1917 (2016). [PubMed: 27137495]
70. Sanders TJ, Marshall CJ & Santangelo TJ The Role of Archaeal Chromatin in Transcription. J. Mol. Biol (2019). doi:10.1016/j.jmb.2019.05.006
71. Dar D et al. Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. Science 352, aad9822 (2016). [PubMed: 27120414]
72. Spitalny P & Thomm M A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. Mol. Microbiol 67, 958–970 (2008). [PubMed: 18182021]
73. Blombach F, Matelska D, Fouqueau T, Cackett G & Werner F Key Concepts and Challenges in Archaeal Transcription. J. Mol. Biol (2019). doi:10.1016/j.jmb.2019.06.020

74. Shashni R, Qayyum MZ, Vishalini V, Dey D & Sen R Redundancy of primary RNA-binding functions of the bacterial transcription terminator Rho. Nucleic Acids Res. 42, 9677–90 (2014). [PubMed: 25081210]

75. Valabhoju V, Agrawal S & Sen R Molecular Basis of NusG-mediated Regulation of Rho-dependent Transcription Termination in Bacteria. J. Biol. Chem 291, 22386–22403 (2016). [PubMed: 27605667]

76. Peters JM et al. Rho and NusG suppress pervasive antisense transcription in Escherichia coli. Genes Dev. 26, 2621–2633 (2012). [PubMed: 23207917]

77. Fong N et al. Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition. Mol. Cell 60, 256–67 (2015). [PubMed: 26474067]

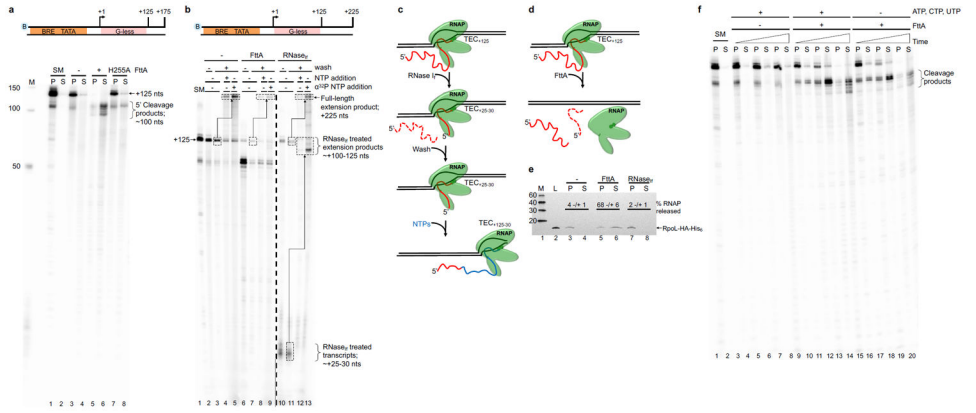


Figure 1. FttA is a *bona fide* archaeal transcription termination factor.

a. Transcripts within intact TECs are retained in pellet (P) fractions; transcripts released from terminated TECs partition into the supernatant (S). Radiolabeled transcripts within starting material (SM) $TECs_{+125}$ and mock treated $TECs_{+125}$ are retained in pellet fractions (lanes 1-4), whereas $FttA^{WT}$ addition results in cleavage of nascent transcripts and termination of most TECs (lanes 5-6). A catalytically deficient FttA variant ($FttA^{H255A}$) abrogates cleavage and RNA release (lanes 7-8). Lane M contains ^{32}P-labeled ssDNA markers. **b.** FttA-mediated termination is distinct from RNase treatment of intact TECs. $TECs_{+125}$ (SM, lane 1) are resistant to repeated washes and readily resume elongation upon NTP addition to generate +225 nt transcripts (lanes 2-5). Dashed boxes and arrows denote +125 transcripts that are elongated to +225 nt transcripts; the specific activity of +225 transcripts can be increased by addition of additional ^{32}P-α-UTP during resumed elongation. RNase I_f digestion of nascent transcripts associated with washed $TECs_{+125}$ results in degradation of the nascent transcript to just ~20-30 nts, but TECs with shortened transcripts remain associated with the DNA and survive repeated washing (lanes 10-11). $TECs_{\sim+25\text{-}30}$ resultant from RNase I_f treatment of $TECs_{+125}$ readily resume elongation upon NTP addition to generate ~+125 nt full-length transcripts (lanes 12-13). Dashed boxes and arrows denote ~+25 transcripts that are elongated to ~+125 nt transcripts; the specific activity of ~+125 transcripts can be increased by addition of ^{32}P-α-UTP during resumed elongation. FttA addition to TECs+125 disrupts most TECs with nascent transcript cleavage (lanes 6-9), results in release of most TECs from the template and cleaved transcripts cannot be extended by NTP addition (lanes 8-9). **c** and **d**. Diagrams of the fate of $TECs_{+125}$ following RNase I_f and FttA treatment, respectively. **e.** FttA releases RNAP from the DNA template into solution confirming dissociation of the TEC and *bona fide* FttA-mediated transcription termination. RNAP is tracked and quantified by Western blots (n = 3 independent replicates) with anti-HA antibodies that recognize the modified RpoL-subunit. **f.** FttA is not reliant on NTP hydrolysis to inactivate TECs, cleave nascent transcripts and terminate transcription. For panels a, b f: Similar results were observed in 4 independent experiments.

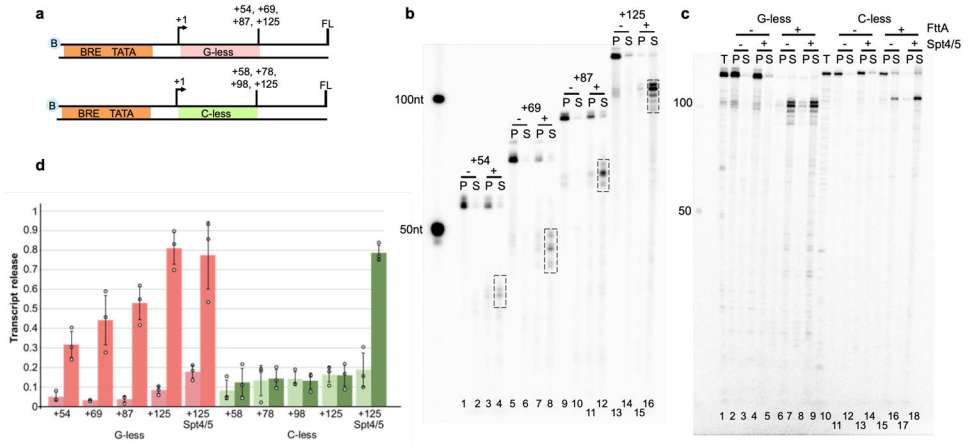


Figure 2. FttA-mediated termination shares mechanistic requirements of rho-mediated bacterial transcription termination.

a. Promoter-directed transcription of biotinylated templates encoding G-less or C-less cassettes permits formation of TECs with increasing length A-, C-, and U-rich, or A-, G-, and U-rich nascent transcripts, respectively. FL = full-length; all templates permit elongation for 100 nts beyond the G- or C-less cassette. **b.** TECs remain stably associated and transcripts are primarily recovered in the pellet (P) fraction in the absence (−) of FttA. When FttA is present (+), transcripts are cleaved and primarily recovered in the supernatant (S) fraction. Cleavage releases ~20 – 30 nt shorter transcripts (boxed). The left-most lane contains ^{32}P-labeled ssDNA markers. **c.** Addition of Spt4-Spt5 largely abrogates the RNA sequence-requirements of FttA-mediated transcription termination. T= total reaction = P+S. The left-most lane contains ^{32}P-labeled ssDNA markers. **d.** Transcript release was quantified with and without FttA addition for TECs with increasing length transcripts on G-less (pink/salmon) and C-less cassettes (mint/green), with and without Spt4-Spt5 addition for TECs $_{+125}$ formed on G- and C-less cassettes. Error bars were calculated as standard deviation from the mean (n = 3 independent experiments). For panels b, c: Similar results were observed in 3 independent experiments.

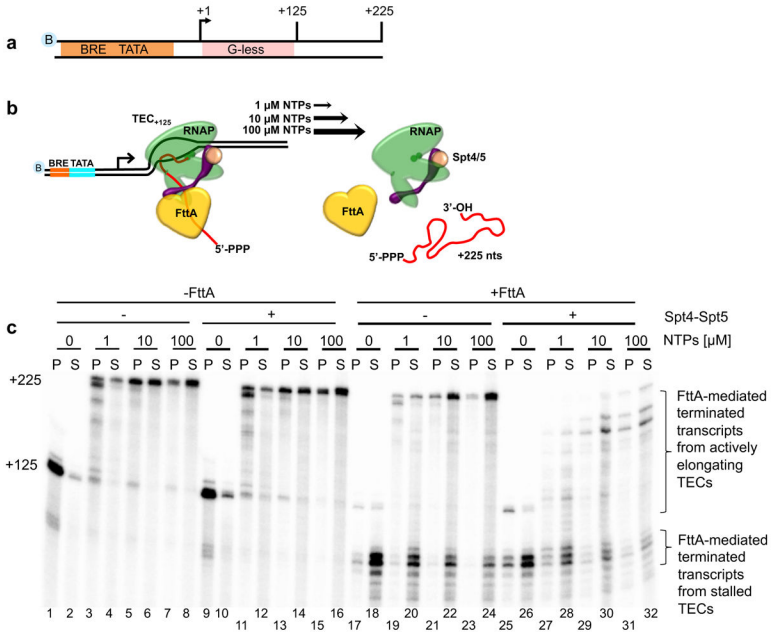


Figure 3. FttA-mediated transcription termination is competitive with transcription elongation.

a and **b**. Washed, NTP-deprived TECs$_{+125}$ were assembled on biotinylated templates with a +125 nt G-less cassette. Resumed elongation upon differential [NTP] addition permits transcription to generate +225 nt transcripts, albeit at different rates. **c**. FttA readily terminates stalled or slowly elongating TECs (lanes 17-24) and FttA-mediated termination becomes competitive with transcription elongation even at high [NTP] in the presence of Spt4-Spt5 (lanes 25-32). Similar results were observed in 3 independent experiments.

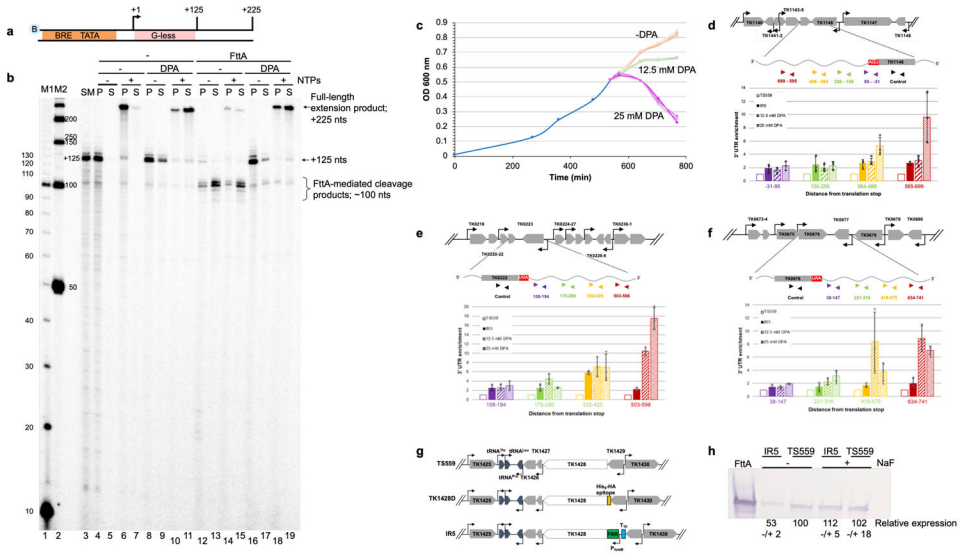


Figure 4. Inhibition of FttA activity abolishes transcription termination *in vitro* and reduced FttA-expression or activity alters steady-state RNA 3'-termini *in vivo*.

a. and **b.** $TECs_{+125}$ (lane 3) resume elongation upon NTP addition to generate +225 nt transcripts −/+ 25 mM DPA (lanes 4-11). FttA addition results in transcript cleavage and release of most TECs to the supernatant - DPA (lanes 12-13), inhibiting resumed elongation upon NTP addition (lanes 14-15). Pre-incubation of FttA with 25 mM DPA inhibits FttA-mediated termination (lanes 16-17), permitting $TECs_{+125}$ to resume elongation (lanes 18-19). Lanes M1 and M2 contain ^{32}P-labeled 10- and 50-nt ssDNA markers, respectively. Similar results were observed in 5 independent experiments. **c.** Inhibition of metallo-beta-lactamase/beta-CASP protein activity impairs growth of *T. kodakarensis.* A mid-log culture of *T. kodakarensis* strain TS559 was split into nine cultures, with three biological replicates exposed to 0 mM (peach), 12.5 mM (green) or 25 mM DPA (purple). **d, e and f**. RNAs recovered one-hour post DPA-addition to cultures of TS559, or from cultures of IR5 grown in the absence of NaF display altered 3'-termini. TRIzol extracted RNAs were reverse transcribed with primers complementary to nascent transcript sequences of TK1146, TK0222 and TK0676 to generate cDNAs that were quantified and normalized to internal controls. Inhibiting FttA-activity with DPA or lowering steady-state FttA levels by riboswitch-mediated controlled expression impacts the abundance of RNAs with extended 3'-termini *in vivo*. RNA abundance in untreated TS559 cultures (open bars) was set to 1.0, and fold changes in the abundance of amplicons reflecting RNA transcripts with extended 3'-sequences at increasing distances from the translation stop site (purple, green, orange and red) are shown for strain IR5 (solid bars), TS559 + 12.5 mM DPA (wide stripes) and TS559 + 25 mM DPA (narrow stripes). Errors were calculated at a 95% confidence interval with the center value as the mean of 3 biological replicates. **g.** Maps of the TK1428 locus in parental (TS559), N-terminally tagged (TK1428D) and riboswitch-regulated expression (IR5) strains. **h.** Western blots demonstrate the reduction in FttA protein levels in strain IR5 upon removal of NaF from the medium; n = 3 independent replicates. Size standards are identified by MW (left).